

The California Spatial Information Library, CaSIL Users Guide – Lilac Distribution

Quinn Hart, Wayland Yu {quinn,wayland.yu}@ceres.ca.gov

October 2003

CaSIL is a Library of Spatial Data products of importance to the State of California, by providing a simple interface to spatial data products, that is accessible to the largest audience possible. This user's guide describes The California Spatial Information Library (CaSIL), what it is, how it is organized, and demonstrates how best to use the library effectively.

Contents

1	CaSIL Overview	2
1.1	New Features	2
2	New Data	3
3	Using CaSIL	4
3.1	CaSIL Organization	4
4	CaSIL Mirroring	5
4.1	CaSIL Mirroring via the Mirror Program	6
4.2	CaSIL Mirroring via Wget	6
5	Architecture Overview	7
6	Upcoming Distribution	7
A	Windows Specific Accessibility	8
A.1	via Internet Explorer	9
A.2	via My Network Places	9
B	Downloading Tiff files via Internet Explorer	10

1 CaSIL Overview

The California Spatial Information Library, or CaSIL, is a Library of Spatial Data products of importance to the State of California. CaSIL provides a simple, intuitive interface to spatial data products, that is accessible to the largest audience possible.

CaSIL can be accessed via the California Environmental Information Catalog (CEIC) at [CEIC Library](#).

The following are items of note regarding CaSIL:

CaSIL as a GIS Filesystem CaSIL acts as a large online GIS filesystem, that hosts datasets from several sources, and hosts the value-added datasets that were derived from the original datasets. By value-added, we mean that the original datasets were reprojected to a canonical projection. At CaSIL, the canonical projection is the Albers Equal Area projection, which allows the overlaying of various datasets with one another.

We also provided additional metadata, such as that which conforms to the FGDC metadata content standard, to help better identify to the users the content of these datasets.

By hosting the original datasets and the reprojected datasets, users can access both types of datasets from one convenient, centralized location.

CaSIL Technology CaSIL does not provide its own integrated web GIS viewer; however, CaSIL does provide several Internet Map Servers, that can be accessed from [CaSIL Internet Map Servers](#).

Two different map servers can be accessed from the above site, which are based on ArcIMS. The first site provides access to the approximately 53 distinct data layers that are available at CaSIL, which are organized by theme. This site can be accessed directly at [Map Server by Theme](#).

The second site provides an interface for downloading DRGs for California, as provided by the USGS. The datasets are available at 1:250K, 1:100K, and 1:24K scale. This site can be accessed directly at [CA DRG viewer](#).

A third interface, where the data layers are organized by data theme, is currently under construction.

1.1 New Features

Many new features are added to the latest distribution of CaSIL (*Lilac*).

The most important feature is the inclusion of more data from various contributors, such as the California government districts (senate, congressional, assembly).

The CaSIL file system is reorganized to reflect the diverse needs of the many types of users who access it. Besides GIS professionals, we anticipate more casual users to access CaSIL, such as

educators from various levels of education. Regardless of access method, the file system structure is organized in such a manner as to facilitate a consistent means of accessing data. Files can be accessed via FTP and HTTP protocols, with others protocols to be supported in the near future. The file system is also reorganized in such a fashion as to enable LAN administrators to more effectively mirror the files contained therein.

Metadata conforming to the FGDC Metadata Content Standard, and distributed in the industry standard XML format, is created and associated with datasets, to help clients better determine the content of the datasets. We have tested most FGDC metadata and have verified their readability with respect to the following systems: ArcCatalog, Explorer, and XMLlint.

2 New Data

The following new data products are now included in the CaSIL distribution:

Digital Ortho Quadrangles The largest amounts of data included in the latest distribution of CaSIL are DOQ data. This data initially comes from USGS, but has been made available by the time and support of [CA Department of Conservation](#), [CA Department of Fish and Game](#), and many other state organizations.

Currently, CaSIL hosts 11K DOQQ datasets, for a total storage usage of 650 GB. Different types of DOQQs are available (Black and White, Color, Quarter or whole quad) based mainly on availability from USGS.

National Hydrology Dataset (NHD) New datasets from the National Hydrology Dataset are incorporated from [The State Water Resources Control Board](#).

Overview Datasets Value-added datasets created at CaSIL include slightly modified National Atlas datasets. Datasets that serve as indices can be found at [CA Indices](#).

US Census, Tiger Data Tiger/LINE data based on the latest US Census Bureau Tiger 2K data, reprojected in the standard CaSIL projection (that is, the Albers Equal Area Projection), can be found at [Tiger2k](#).

Political Boundaries The latest California Senate, Congressional, and Assembly districts, which are reprojected to the standard CaSIL projection, can be found at [CA Districts](#).

DRGs The latest distribution of CaSIL includes the original DRGs that were based upon UTM projection, and distributed in the MrSID (Multi-resolution Seamless Image Database) compression format. These datasets are reprojected to the CaSIL standard projection, and organized based on series (C, F, O), and datum (NAD27, NAD83, etc.) A complete description of the DRGs hosted by CaSIL, as well as the dedicated website at CaSIL for DRG images, can be found in the section CaSIL Organization, below.

SPOT CaSIL also hosts SPOT data, based on data collected by the SPOT satellite system and operated by the French Space Agency, le Centre National d'Etudes Spatiales (CNES). Licensing restrictions prevent this dataset from being available directly from the CaSIL library.

TM [TM Data](#), or Thematic Mapper data, are also included at CaSIL.

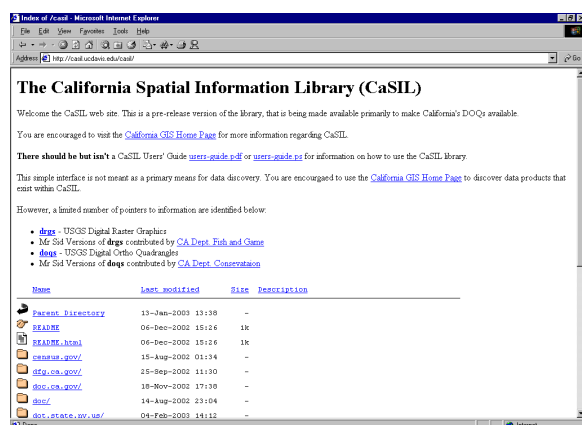
Locally contributed data Data from city and county governments are also hosted at CaSIL, such as the Kern COG datasets that cover the County of Kern and the eleven incorporated cities within Kern County. Datasets from Kern COG are also included at CaSIL at [KernCog datasets](#).

3 Using CaSIL

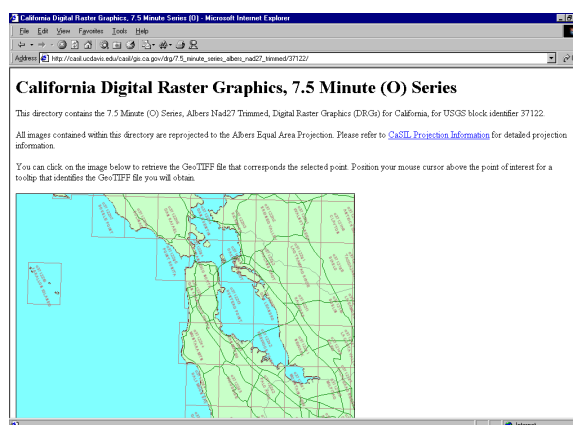
For direct browser access to CaSIL, point your browser to [CaSIL website](#). For FTP access, please access [CaSIL FTP](#).

3.1 CaSIL Organization

Figure 1, panel (a), shows what you will see when you access the main CaSIL page via the HTTP protocol.



(a)



(b)

Figure 1: CaSIL WebSite

As can be seen here, CaSIL not only serves as a repository for GIS data pertaining to California, but it also reprojects many such datasets to the Albers Equal Area Projection, and deposits these datasets in the directory [gis.ca.gov](#).

For example, the original senate district datasets can be found at [Original Senate Districts](#), whereas

the datasets that were reprojected to the Albers projection by CaSIL is located at [CaSIL Reprojected Senate Districts](#).

We house both the original datasets, as well as the datasets reprojected by CaSIL, to provide users with a common destination for accessing one, or both types of datasets.

For the current distribution of CaSIL, specifically for the DRG section of CaSIL, we will provide client-side image maps to allow users to select datasets by an intuitive map interface. For example, suppose a user wishes to access a 7.5 minute (O) DRG series dataset for California, perhaps the San Francisco Bay Area. The URL for this dataset, which corresponds to block 37122, is at [DRG O Series](#). When a user traverses to the following directory, the user will immediately see the following client-side image map, as shown in figure 1, panel (b).

The client-side image map is currently configured so that, when the user clicks on a specific point on the graphic, the user will retrieve a GeoTIFF file that corresponds to the selected point.

For users accessing the DRG directories via HTTP, we have also included brief descriptions at the bottom of each page, next to each file name, to help users determine at a glance exactly what region a specific file covers; for example:

```
o37122g4.tif      17-Dec-2002 10:24      1k  SAN_FRANCISCO_NORTH
```

Eventually, this point and click interface, provided by the client-side image map, will be extended to the various DOQQ directories as well.

4 CaSIL Mirroring

CaSIL has been redesigned to be more convenient for mirroring by redistribution sites, and for local LAN administrators. With its exposed filesystem structure, users can maintain large sections of CaSIL locally with a minimum amount of effort. The CaSIL ftp site is more reliable for maintaining consistent copies of CaSIL data, but mirroring via the http site is also possible, but involves some extra effort.

Be forewarned that you can end up mirroring a lot of data if you are not sure of how to use your mirror program of choice, since CaSIL contains many large datasets.

In the following examples, we will assume that the user will construct a local mirror copy in `/opt/casil/` for Un*x systems, or `p:/casil` for Windows systems.

The following sections describe how to mirror the CaSIL repository via two common mirroring programs: *Mirror* and *Wget*.

4.1 CaSIL Mirroring via the Mirror Program

Here we give an example of an FTP mirroring application with *Mirror*; note however that any FTP mirroring program can be used instead. *Mirror* can easily be used to copy subdirectories at CaSIL. *Mirror* is written in Perl, and runs on Unix/Linux, as well as Microsoft Windows platforms. You can retrieve the latest version from <http://sunsite.org.uk/packages/mirror/>.

For example, to construct a mirror of the Teale datasets in gis.ca.gov, specify the following:

```
mirror -g casil.ucdavis.edu:/casil/gis.ca.gov/teale \  
      -k local_dir=p:/casil/gis.ca.gov/teale
```

Here, the site of the mirrored copy, as well as the pathname to the directory where the mirror is constructed from, is specified using the **-g** parameter. The **-k** parameter assigns a value to a variable; here, the *local_dir* variable is assigned the name of the directory where the mirror is constructed, in this case “p:/casil/gis.ca.gov/teale”. In this example, the value given is an MS-DOS style directory, and the path as a whole is consistent with the file structure at CaSIL.

Alternatively, you can specify package descriptions for mirroring certain sections of CaSIL. Example package descriptions are shown in <http://casil.ucdavis.edu/casil/support/mirror/packages/casil>. See <http://sunsite.org.uk/packages/mirror/mirror.html> for information about the format of the package description file.

For example, to use a package file to generate a mirror of the Teale datasets, first construct a package file, by entering the following lines in the file:

```
package=teale  
comment=Teale Datasets at gis.ca.gov  
site=casil.ucdavis.edu  
remote_dir=/casil  
get_patt=gis.ca.gov/teale  
local_dir=p:/casil/gis.ca.gov/teale
```

If the above package file is called p:/casil/support/mirror/packages/casil, you can then construct your mirror with the following command:

```
mirror -p teale p:/casil/support/mirror/packages/casil
```

The selected package to mirror (in this case, *teale*) is specified using the **-p** option.

4.2 CaSIL Mirroring via Wget

Wget, at <http://www.gnu.org/software/wget/wget.html>, is a common file retrieval/mirroring tool that is standard on most Unix/Linux systems. A Windows port is available with the freely available

Cygwin distribution. A comprehensive manual on *Wget* is available at <http://www.gnu.org/manual/wget-1.8.1/>.

For example, to copy all files and subdirectories from Teale using *Wget*, run the following command from within the “/opt” directory:

```
wget -m -np -nH -R ‘\?’ \
http://casil.ucdavis.edu/casil/gis.ca.gov/teale -P /opt
```

The **-m** option tells *Wget* to enter mirroring mode, so that files are downloaded recursively. Mirror mode also activates timestamping, so that only recently updated files at CaSIL are downloaded. The **-R** option prevents the download of files that match the specified pattern; here the pattern prevents files of the form “?D=A” from being downloaded. The **-np** option instructs the program to never go up from the parent directory, which is necessary in conjunction with the mirroring option. The **-nH** option prevents a “casil.ucdavis.edu” parent directory from being created.

The **-P** option specifies the directory prefix; this specifies the local directory where the mirror should be constructed. In the above example, the directory “casil/gis.ca.gov/teale” on CaSIL is mirrored to the local “/opt” directory.

An even better strategy is demonstrated in the following example: this example allows the user to run *Wget* from ANY local directory, but still have the mirror copy constructed at the right location.

```
wget -m -np -nH -R ‘\?’ --cut-dirs=1 \
http://casil.ucdavis.edu/casil/gis.ca.gov/teale -P /opt/casil
```

The **-cut-dirs** option specifies how many parent directory components to eliminate. A value of 1 leaves you with “gis.ca.gov/teale”; this directory is then mirrored at the local “/opt/casil” directory.

5 Architecture Overview

Figure 2 provides the user with a high-level view of the CaSIL architecture. From this figure, it can be seen how datasets from various contributors, such as the [U.S. Census Bureau](#) and the [California Department of Fish and Game](#), are collected at CaSIL, serving as the centralized dataset repository. The [CaSIL website](#) then serves the datasets as originally provided by our contributors. Datasets generated by CaSIL can be found at [CaSIL datasets](#). In particular, DRGs and DOQQs reprojected by CaSIL can be found at the latter location.

6 Upcoming Distribution

For the next distribution, we intend to address the following issues.

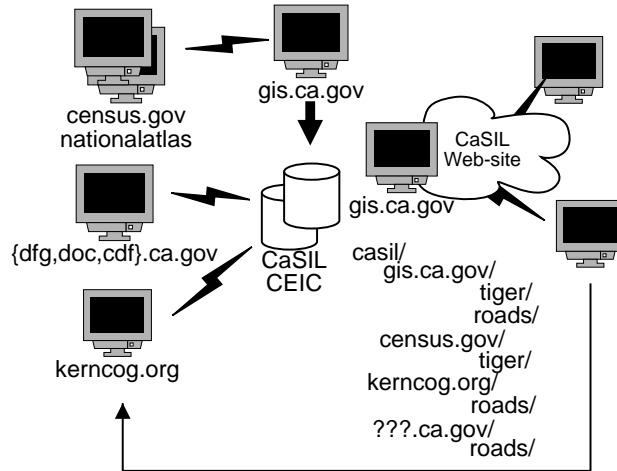


Figure 2: CaSIL Architecture

To improve quality of service (QoS) for our users, we will introduce additional, more reliable mirrors to better service regional requests. Current FTP access statistics at CaSIL indicate a mean throughput of 5GB per day, which translates to a 1.3TB throughput for a given year. This figure underscores the need for better mirrors to provide for the best bandwidth utilization possible.

Most of our datasets, such as the DRGs, currently conform to the NAD27 datum; the next distribution of CaSIL will provide such datasets in the NAD83 datum as well.

We intend to provide better attribute support in our metadata for the next distribution. Two areas where we can improve our attribute support will come from the [GML](#) and [Dublin Core](#) (ISO/IEC 11179 [ISO11179]) standards.

To facilitate interoperability, the next distribution will have extended OpenGIS functionality. Specifically, CaSIL will provide access methods that conform to the industry standard [OpenGIS Web Map Service](#) and [OpenGIS Web Feature Service](#) specifications.

The next distribution will have several additions to the CEIC. Extended attribution information will be incorporated into the metadata for the datasets. More distribution options will be provided, in addition to the standard FTP and HTTP file access mechanisms. Metadata ingest will also be incorporated.

Lastly, for users who wish to retrieve the latest, bleeding edge releases (hyperbole) of our datasets, and for those who wish to contribute to CaSIL, we will provide Version Control access to the CaSIL repository via the CVS mechanism.

A Windows Specific Accessibility

Users who access CaSIL via FTP using a web page browser, such as Microsoft's Internet Explorer or Netscape Navigator, traditionally browse the contents at CaSIL by clicking on hyperlinks. However,

A.1 via Internet Explorer

Figure 1 consists of two side-by-side screenshots of the 'Internet Options' dialog box in Internet Explorer. The left screenshot shows the 'Content Advisor' tab, where the 'Content Advisor' is enabled and set to 'Moderate'. The right screenshot shows the 'Content Advisor' tab, where the 'Content Advisor' is disabled.

Under the “Browsing” header, there is an option “Enable folder view for FTP sites”. To browse FTP sites such as CaSIL using a Windows Explorer style interface, enable the check box next to this option. Once you do this, you can open the CaSIL FTP site at “ftp://casil.ucdavis.edu/casil”. Figure 3, panel (b) shows how the CaSIL FTP site looks like from Internet Explorer, with Windows Explorer style browsing enabled.

The second option is somewhat more complicated to set up, but allows the user to access her favorite network sites (including FTP sites) from one convenient, centralized location, by accessing the “My Network Places” folder. Note that this folder is only present in Windows 2000 and Windows XP;

please refer to the previous section for users of Windows 9x or Windows NT 4.0 to access CaSIL using folder view in Internet Explorer.

To access CaSIL via the My Network Places folder, first select the folder. This folder is illustrated in figure 4, panel (a). Double click the “Add Network Place” icon in the folder. Panel (b), shows the “Add Network Place Wizard” dialog. Click the next button. The next dialog, panel (c), allows you to choose a service provider. For our purposes, select the “Choose another network location”, and click the next button; this option allows us to specify an FTP site. The next dialog, panel (d), gives the user a place to enter the desired network (in this case, FTP) location; enter “ftp://casil.ucdavis.edu/casil” in the text box and click the next button. The next dialog, panel (e), allows the user to choose whether to log into the FTP site anonymously. Most users will choose anonymous access, so click on the “Log on anonymously” check box, and click the next button. The next dialog, panel (f), lets the user assign a name to the FTP site; here, we have named the site “CaSIL FTP site”. The final dialog, panel (g), informs the user that the CaSIL FTP link is successfully established in the “My Network Places” folder. If any of the information from this dialog box is incorrect, hit the back button and make the appropriate changes. Select the finish button to complete the setup. Assuming the setup is successful, panel (a), shows the CaSIL FTP site in the “My Network Places” folder, and panel (h) shows the CaSIL FTP site in folder view.

B Downloading Tiff files via Internet Explorer

Users of Internet Explorer who access Tiff files are advised to not click on the link to the image, but should instead save the files directly, by right clicking the link, and choose “Save Target As” from the context menu. The user can then view the images properly in their software of choice, such as ESRI ArcExplorer or Adobe Photoshop.

There is a well known issue in Internet Explorer regarding its handling of Tiff images, in that some images are displayed as garbled text when one clicks on the link. The issue does not appear in alternate browsers such as Opera, Mozilla, or Netscape 7.

The following thread provides additional information about this issue:

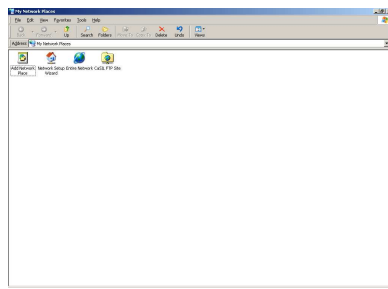
<http://remotesensing.org/geotiff/mail/msg00704.html>

At CaSIL, we highly recommend against doing the following, but if you feel that Internet Explorer is the root of the problem, and are confident in using Registry Editor, you might want to try the steps in the following two Microsoft Knowledge Base articles (in particular the first one):

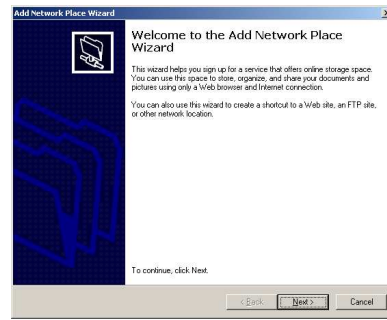
<http://support.microsoft.com/?kbid=319829>

<http://support.microsoft.com/?kbid=320289>

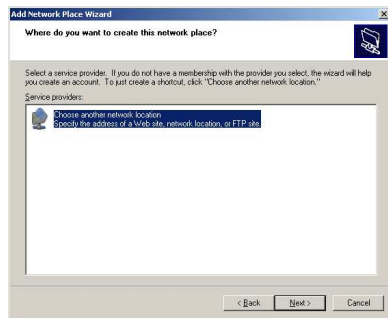
They both involve using the Registry Editor, so if you’re not sure what you’re doing, you could seriously damage your operating system in the process. Most likely, the easiest way to download the file of interest is to use one of the alternative browsers listed for the download.



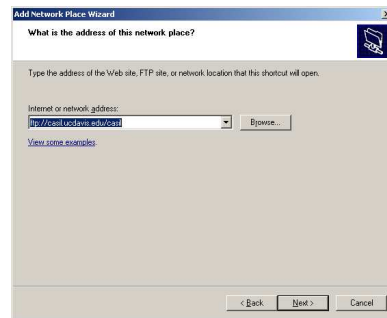
(a)



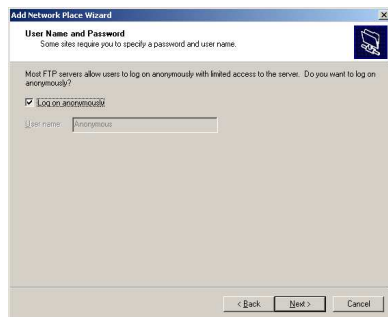
(b)



(c)



(d)



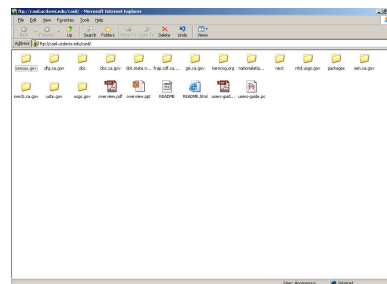
(e)



(f)



(g)



(h)

Figure 4: My Network Places Setup